

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

(i) APPLICANT: Bandman, Olga  
Hawkins, Phillip R.

(ii) TITLE OF THE INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN

(iii) NUMBER OF SEQUENCES: 5

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
(B) STREET: 3174 Porter Drive  
(C) CITY: Palo Alto  
(D) STATE: CA  
(E) COUNTRY: US  
(F) ZIP: 94304

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ Version 1.5

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Filed Herewith

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/088,641  
(B) FILING DATE: June 2, 1998

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.  
(B) REGISTRATION NUMBER: 36,749  
(C) REFERENCE/DOCKET NUMBER: PF-0163 US

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415-855-0555  
(B) TELEFAX: 415-845-4166

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 472 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: None

## (vii) IMMEDIATE SOURCE:

(A) LIBRARY:  
(B) CLONE: 989953

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Ala	Thr	Lys	Cys	Gly	Asn	Cys	Gly	Pro	Gly	Tyr	Ser	Thr	Pro	Leu
1				5				10						15	
Glu	Ala	Met	Lys	Gly	Pro	Arg	Glu	Glu	Ile	Val	Tyr	Leu	Pro	Cys	Ile
			20					25					30		
Tyr	Arg	Asn	Thr	Gly	Thr	Glu	Ala	Pro	Asp	Tyr	Leu	Ala	Thr	Val	Asp
		35					40					45			
Val	Asp	Pro	Lys	Ser	Pro	Gln	Tyr	Cys	Gln	Val	Ile	His	Arg	Leu	Pro
50						55					60				

F0163-2 DIV

THE  
FEDERAL  
BUREAU OF  
INVESTIGATION  
OF THE  
DEPARTMENT OF JUSTICE  
WASHINGTON, D. C. 20535

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1711 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: None

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY:  
 (B) CLONE: 989953

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGGCCTCTG ACACCAGCAC AGCAAACCCG CCGGGATCAA AGTGTACCAG TCGGCAGCAT 60  
 GGCTACGAAA TGTGGGAATT GTGGACCCGG CTACTCCACC CCTCTGGAGG CCATGAAAGG 120  
 ACCCAGGGAA GAGATCGTCT ACCTGCCCTG CATTTACCGA AACACAGGCA CTGAGGCCCC 180  
 AGATTATCTG GCCACTGTGG ATGTTGACCC CAAGTCTCCC CAGTATTGCC AGGTCATCCA 240  
 CCGGCTGCCC ATGCCCAACC TGAAGGACGA GCTGCATCAC TCAGGATGGA ACACCTGCAG 300  
 CAGCTGCTTC GGTGATAGCA CCAAGTCGCG CACCAAGCTG GTGCTGCCCC GTCTCATCTC 360  
 CTCTCGCATC TATGTGGTGG ACGTGGGCTC TGAGCCCCGG GCCCCAAAGC TGCACAAGGT 420  
 CATTGAGCCC AAGGACATCC ATGCCAAGTG CGAACTGGCC TTTCTCCACA CCAGCCACTG 480  
 CCTGGCCAGC GGGGAAGTGA TGATCAGCTC CCTGGGAGAC GTCAAGGGCA ATGGCAAAGG 540  
 GGGTTTTGTG CTGCTGGATG GGGAGACGTT CGAGGTGAAG GGGACATGGG AGAGACCTGG 600  
 GGGTGCTGCA CCGTTGGGCT ATGACTTCTG GTACCAGCCT CGACACAATG TCATGATCAG 660  
 CACTGAGTGG GCAGCTCCCA ATGTCTTACG AGATGGCTTC AACCCCGCTG ATGTGGAGGC 720  
 TGGACTGTAC GGGAGCCACT TATATGTATG GGACTGGCAG CGCCATGAGA TTGTGCAGAC 780  
 CCTGTCTCTA AAAGATGGGC TTATTCCTTT GGAGATCCGC TTCCTGCACA ACCCAGACGC 840  
 TGCCCAAGGC TTTGTGGGCT GCGCACTCAG CTCCACCATC CAGCGCTTCT ACAAGAACGA 900  
 GGGAGGTACA TGGTCACTGG AGAAGGTGAT CCAGGTGCCC CCAAGAAAG TGAAGGGCTG 960  
 GCTGCTGCCC GAAATGCCAG GCCTGATCAC CGACATCCTG CTCTCCCTGG ACGACCGCTT 1020  
 CCTTACTTTC AGCAACTGGC TGCATGGGGA CCTGAGGCAG TATGACATCT CTGACCCACA 1080  
 GAGACCCCGC CTCACAGGAC AGCTCTTCCT CGGAGGCAGC ATTGTTAAGG GAGGCCCTGT 1140  
 GCAAGTGCTG GAGGACGAGG AACTAAAGTC CCAGCCAGAG CCCCTAGTGG TCAAGGGAAA 1200  
 ACGGGTGGCT GGAGGCCCTC AGATGATCCA GCTCAGCCTG GATGGGAAGC GCCTCTACAT 1260  
 CACCACGTCG CTGTACAGTG CCTGGGACAA GCAGTTTAC CCTGATCTCA TCAGGGAAGG 1320  
 CTCTGTGATG CTGCAGGTTG ATGTAGACAC AGTAAAAGGA GGGCTGAAGT TGAACCCCAA 1380  
 CTTCTGCTG GACTTCGGGA AGGAGCCCTT TGGCCCAGCC CTTGCCCATG AGCTCCGCTA 1440  
 CCCTGGGGGC GATTGTAGCT CTGACATCTG GATTTGAACT CCACCCTCAT CACCCACACT 1500  
 CCCTATTTTG GGCCCTCACT TCCTTGGGGA CCTGGCTTCA TTCTGCTCTC TCTTGGCACC 1560  
 CGACCCTTGG CAGCATGTAC CACACAGCCA AGCTGAGACT GTGGCAATGT GTTGAATCAT 1620  
 ATACATTTAC TGACCACTGT TGCTTGTTCG TCACTGTGCT GCTTTTCCAT GAGCTCTTGG 1680  
 AGGCACCAAG AAATAAACTC GTAACCCTGT C 1711

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: None

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank  
 (B) CLONE: 1374792

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ala Thr Lys Cys Gly Asn Cys Gly Pro Gly Tyr Ser Thr Pro Leu  
 1 5 10 15  
 Glu Ala Met Lys Gly Pro Arg Glu Glu Ile Val Tyr Leu Pro Cys Ile  
 20 25 30

Tyr Arg Asn Thr Gly Thr Glu Ala Pro Asp Tyr Leu Ala Thr Val Asp  
 35 40 45  
 Val Asp Pro Lys Ser Pro Gln Tyr Cys Gln Val Ile His Arg Leu Pro  
 50 55 60  
 Met Pro Asn Leu Lys Asp Glu Leu His His Ser Gly Trp Asn Thr Tyr  
 65 70 75 80  
 Ser Ser Cys Phe Gly Asp Ser Thr Lys Ser Arg Asn Lys Leu Val Leu  
 85 90 95  
 Pro Ser Leu Ile Ser Ser Arg Ile Tyr Val Val Asp Val Gly Ser Glu  
 100 105 110  
 Pro Gly Pro Gln Lys Leu His Lys Val Ile Glu Pro Lys Asp Ile His  
 115 120 125  
 Ala Lys Cys Glu Leu Ala Cys Leu His Thr Ser His Cys Leu Ala Ser  
 130 135 140  
 Gly Glu Val Met Ile Ser Ser Leu Gly Asp Val Lys Gly Asn Gly Lys  
 145 150 155 160  
 Gly Gly Phe Val Leu Leu Asp Gly Glu Thr Phe Glu Val Lys Gly Thr  
 165 170 175  
 Trp Glu Arg Pro Gly Gly Ala Ala Pro Leu Gly Tyr Asp Phe Trp Tyr  
 180 185 190  
 Gln Pro Arg His Asn Val Met Ile Ser Thr Glu Trp Ala Ala Pro Asn  
 195 200 205  
 Val Leu Arg Asp Gly Phe Asn Pro Ala Asp Val Glu Ala Gly Leu Tyr  
 210 215 220  
 Gly Ser His Leu Tyr Val Trp Asp Trp Gln Arg His Glu Ile Val Gln  
 225 230 235 240  
 Thr Leu Ser Leu Lys Asp Gly Leu Ile Pro Leu Glu Ile Arg Phe Leu  
 245 250 255  
 His Asn Pro Ser Ala Thr Gln Gly Phe Val Gly Cys Ala Ser Ala Pro  
 260 265 270  
 Asn Ile Gln Arg Phe Tyr Lys Thr Arg Glu Gly Thr Trp Ser Val Glu  
 275 280 285  
 Lys Val Ile Gln Val Pro Pro Lys Lys Val Lys Gly Trp Leu Leu Pro  
 290 295 300  
 Gly Val Pro Gly Leu Ile Thr Asp Ile Leu Leu Ser Leu Asp Asp Arg  
 305 310 315 320  
 Phe Leu Tyr Phe Ser Asn Trp Leu His Gly Asp Leu Arg Gln Tyr Asp  
 325 330 335  
 Ile Ser Asp Pro Gln Arg Pro Arg Leu Thr Gly Gln Leu Phe Leu Gly  
 340 345 350  
 Gly Ser Ile Val Lys Gly Gly Pro Val Gln Val Leu Glu Asp Glu Glu  
 355 360 365  
 Leu Lys Ser Gln Pro Glu Pro Leu Val Val Lys Gly Lys Arg Val Ala  
 370 375 380  
 Gly Gly Pro Gln Met Ile Gln Leu Ser Leu Asp Gly Lys Arg Leu Tyr  
 385 390 395 400  
 Ile Thr Thr Ser Leu Tyr Ser Ala Trp Glu Lys Gln Phe Tyr Pro Asp  
 405 410 415  
 Leu Ile Arg Glu Gly Ser Val Met Leu Gln Val Asp Val Asp Thr Val  
 420 425 430  
 Lys Gly Gly Leu Lys Leu Asn Pro Asn Cys Leu Val Asp Phe Gly Lys  
 435 440 445  
 Glu Pro Leu Gly Pro Ala Leu Ala His Glu Leu Arg Tyr Pro Gly Gly  
 450 455 460  
 Asp Cys Ser Ser Asp Ile Trp Ile  
 465 470

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:  
 (A) LIBRARY: GenBank  
 (B) CLONE: 227630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Ala	Thr	Lys	Cys	Thr	Lys	Cys	Gly	Pro	Gly	Tyr	Ser	Thr	Pro	Leu
1				5					10					15	
Glu	Ala	Met	Lys	Gly	Pro	Arg	Glu	Glu	Ile	Val	Tyr	Leu	Pro	Cys	Ile
			20					25					30		
Tyr	Arg	Asn	Thr	Gly	Thr	Glu	Ala	Pro	Asp	Tyr	Leu	Ala	Thr	Val	Asp
		35					40					45			
Val	Asp	Pro	Lys	Ser	Pro	Gln	Tyr	Ser	Gln	Val	Ile	His	Arg	Leu	Pro
		50				55					60				
Met	Pro	Tyr	Leu	Lys	Asp	Glu	Leu	His	His	Ser	Gly	Trp	Asn	Thr	Cys
65					70					75					80
Ser	Ser	Cys	Phe	Gly	Asp	Ser	Thr	Lys	Ser	Arg	Asn	Lys	Leu	Ile	Leu
			85						90					95	
Pro	Gly	Leu	Ile	Ser	Ser	Arg	Ile	Tyr	Val	Val	Asp	Val	Gly	Ser	Glu
			100					105					110		
Pro	Arg	Ala	Pro	Lys	Leu	His	Lys	Val	Ile	Glu	Ala	Ser	Glu	Ile	Gln
		115					120					125			
Ala	Lys	Cys	Asn	Val	Ser	Ser	Leu	His	Thr	Ser	His	Cys	Leu	Ala	Ser
		130					135					140			
Gly	Glu	Val	Met	Val	Ser	Thr	Leu	Gly	Asp	Ile	Gln	Gly	Asn	Gly	Lys
145						150				155					160
Gly	Ser	Phe	Val	Leu	Leu	Asp	Gly	Glu	Thr	Phe	Glu	Val	Lys	Gly	Thr
			165					170						175	
Trp	Glu	Lys	Pro	Gly	Asp	Ala	Ala	Pro	Met	Gly	Tyr	Asp	Phe	Trp	Tyr
			180					185					190		
Gln	Pro	Arg	His	Asn	Val	Met	Val	Ser	Thr	Glu	Trp	Ala	Ala	Pro	Asn
		195					200					205			
Val	Phe	Lys	Asp	Gly	Thr	Asn	Pro	Ala	His	Val	Glu	Ala	Gly	Leu	Tyr
		210				215					220				
Gly	Ser	Arg	Ile	Phe	Val	Trp	Asp	Trp	Gln	Arg	His	Glu	Ile	Ile	Gln
225					230					235					240
Thr	Leu	Gln	Met	Thr	Asp	Gly	Leu	Ile	Pro	Leu	Glu	Ile	Arg	Phe	Leu
			245						250					255	
His	Asp	Pro	Ser	Ala	Thr	Gln	Gly	Phe	Val	Gly	Cys	Ala	Ser	Ala	Pro
			260					265					270		
Asn	Ile	Gln	Arg	Phe	Tyr	Lys	Asn	Ala	Glu	Gly	Thr	Trp	Ser	Val	Glu
		275					280					285			
Lys	Val	Ile	Gln	Val	Pro	Ser	Lys	Lys	Val	Lys	Gly	Trp	Met	Leu	Pro
		290					295				300				
Gly	Val	Pro	Gly	Leu	Ile	Thr	Asp	Ile	Leu	Leu	Ser	Leu	Asp	Asp	Arg
305					310					315					320
Phe	Leu	Tyr	Phe	Ser	Asn	Trp	Leu	His	Gly	Asp	Ile	Arg	Gln	Tyr	Asp
			325						330					335	
Ile	Ser	Asn	Pro	Gln	Lys	Pro	Arg	Leu	Ala	Gly	Gln	Ile	Phe	Leu	Gly
			340					345					350		

PF-0163-2 DIV

Gly	Ser	Ile	Val	Arg	Gly	Gly	Ser	Val	Gln	Val	Leu	Glu	Asp	Gln	Glu
		355					360					365			
Leu	Thr	Cys	Gln	Pro	Glu	Pro	Leu	Val	Val	Lys	Gly	Lys	Arg	Ile	Pro
	370					375					380				
Gly	Gly	Pro	Gln	Met	Ile	Gln	Leu	Ser	Leu	Asp	Gly	Lys	Arg	Leu	Tyr
385					390					395					400
Ala	Thr	Thr	Ser	Leu	Tyr	Ser	Ala	Trp	Asp	Lys	Gln	Phe	Tyr	Pro	Asp
				405					410					415	
Leu	Ile	Arg	Glu	Gly	Ser	Met	Met	Leu	Gln	Ile	Asp	Val	Asp	Thr	Val
			420					425					430		
Asn	Gly	Gly	Leu	Lys	Leu	Asn	Pro	Asn	Phe	Leu	Val	Asp	Phe	Gly	Lys
	435						440					445			
Leu	Pro	Leu	Gly	Ala	Ala	Leu	Ala	His	Glu	Leu	Arg	Tyr	Pro	Gly	Gly
	450					455					460				
Asp	Cys	Ser	Ser	Asp	Ile	Trp	Ile								
465					470										

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 298710

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Thr	Lys	Cys	Thr	Lys	Cys	Gly	Pro	Gly	Pro	Ser	Thr	Pro	Leu
1				5					10					15	
Glu	Ala	Met	Lys	Gly	Pro	Arg	Glu	Glu	Ile	Val	Tyr	Leu	Pro	Cys	Ile
			20					25					30		
Tyr	Arg	Asn	Thr	Gly	Thr	Glu	Ala	Pro	Asp	Tyr	Leu	Ala	Thr	Val	Asp
		35					40					45			
Val	Asp	Pro	Lys	Ser	Pro	Gln	Tyr	Ser	Gln	Val	Ile	His	Arg	Leu	Pro
	50					55					60				
Met	Pro	Tyr	Leu	Lys	Asp	Glu	Leu	His	His	Ser	Gly	Trp	Asn	Thr	Cys
65					70					75					80
Ser	Ser	Cys	Phe	Gly	Asp	Ser	Thr	Lys	Ser	Arg	Asn	Lys	Leu	Ile	Leu
				85					90					95	
Pro	Gly	Leu	Met	Ser	Ser	Arg	Ile	Tyr	Val	Val	Asp	Val	Gly	Ser	Glu
			100					105					110		
Pro	Arg	Ala	Pro	Lys	Leu	His	Lys	Val	Ile	Glu	Ala	Ser	Glu	Ile	Gln
		115					120					125			
Ala	Lys	Cys	Asn	Val	Ser	Asn	Thr	His	Thr	Ser	His	Cys	Leu	Ala	Ser
	130					135					140				
Gly	Glu	Val	Met	Val	Ser	Thr	Leu	Gly	Asp	Leu	Gln	Gly	Asn	Gly	Lys
145					150					155					160
Gly	Ser	Phe	Val	Leu	Leu	Asp	Gly	Glu	Thr	Phe	Glu	Val	Lys	Gly	Thr
				165					170					175	
Trp	Glu	Lys	Pro	Gly	Gly	Ala	Ser	Pro	Met	Gly	Tyr	Asp	Phe	Trp	Tyr
			180					185					190		
Gln	Pro	Arg	His	Asn	Val	Met	Val	Ser	Thr	Glu	Trp	Ala	Ala	Pro	Asn
		195				200						205			
Val	Phe	Lys	Asp	Gly	Phe	Asn	Pro	Ala	His	Val	Glu	Ala	Gly	Leu	Tyr
	210					215					220				
Gly	Ser	Arg	Ile	Phe	Val	Trp	Asp	Trp	Gln	Arg	His	Glu	Ile	Ile	Gln
225					230					235					240
Thr	Leu	Gln	Met	Thr	Asp	Gly	Leu	Ile	Pro	Leu	Glu	Ile	Arg	Phe	Leu
				245					250					255	
His	Asp	Pro	Ser	Ala	Thr	Gln	Gly	Phe	Val	Gly	Cys	Ala	Leu	Ser	Ser

			260					265					270				
Asn	Ile	Gln	Arg	Phe	Tyr	Lys	Asn	Gly	Glu	Gly	Thr	Trp	Ser	Val	Glu		
		275					280					285					
Lys	Val	Ile	Gln	Val	Pro	Ser	Lys	Lys	Val	Lys	Gly	Trp	Met	Leu	Pro		
	290					295					300						
Glu	Met	Pro	Gly	Leu	Ile	Thr	Asp	Ile	Leu	Leu	Ser	Leu	Asp	Asp	Arg		
305				310						315					320		
Phe	Leu	Tyr	Phe	Ser	Asn	Trp	Leu	His	Gly	Asp	Ile	Arg	Gln	Tyr	Asp		
			325						330					335			
Ile	Ser	Asn	Pro	Gln	Lys	Pro	Arg	Leu	Thr	Gly	Gln	Ile	Phe	Leu	Gly		
		340						345					350				
Gly	Ser	Ile	Val	Arg	Gly	Gly	Ser	Val	Gln	Val	Leu	Glu	Asp	Gln	Glu		
	355						360					365					
Leu	Thr	Cys	Gln	Pro	Glu	Pro	Leu	Val	Val	Lys	Gly	Lys	Arg	Ile	Pro		
	370					375					380						
Gly	Gly	Pro	Gln	Met	Ile	Gln	Leu	Ser	Leu	Asp	Gly	Lys	Arg	Leu	Tyr		
385				390						395					400		
Ala	Thr	Thr	Ser	Leu	Tyr	Ser	Ala	Trp	Asp	Lys	Gln	Phe	Tyr	Pro	Asp		
			405					410						415			
Leu	Ile	Arg	Glu	Gly	Ser	Val	Met	Leu	Gln	Val	Asp	Val	Asp	Thr	Val		
		420						425					430				
Asn	Gly	Gly	Leu	Lys	Leu	Asn	Pro	Asn	Phe	Leu	Val	Asp	Phe	Gly	Lys		
	435					440						445					
Glu	Pro	Leu	Gly	Pro	Ala	Leu	Ala	His	Glu	Leu	Arg	Tyr	Pro	Gly	Gly		
	450					455					460						
Asp	Cys	Ser	Ser	Asp	Ile	Trp	Ile										
465					470												

T0h2h0"892F4B60